



One example of general unidentifiable tensors

Luca Chiantini¹, Massimiliano Mella², Giorgio Ottaviani^{3,*}

¹ *Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche Università di Siena, Pian dei Mantellini 44, 53100 Siena, Italia*

² *Dipartimento di Matematica e Informatica, Università di Ferrara, Via Machiavelli 35, 44121 Ferrara, Italia*

³ *Dipartimento di Matematica e Informatica 'Ulisse Dini' Università di Firenze, Viale Morgagni 67/A, 50134 Firenze, Italia*

Abstract. The identifiability of parameters in a probabilistic model is a crucial notion in statistical inference. We prove that a general tensor of rank 8 in $\mathbb{C}^3 \otimes \mathbb{C}^6 \otimes \mathbb{C}^6$ has at least 6 decompositions as sum of simple tensors, so it is not 8-identifiable. This is the highest known example of balanced tensors of dimension 3, which are not k -identifiable, when k is smaller than the generic rank.

2000 Mathematics Subject Classifications: 14N05, 15A69

Key Words and Phrases: tensor decomposition, identifiability, Segre variety

1. Introduction

The decomposition of tensors $T \in \mathbb{C}^{a_1+1} \otimes \dots \otimes \mathbb{C}^{a_q+1}$ as a sum of *simple* tensors (i.e. tensors of rank 1) is a central problem for many applications of Multilinear Algebra to Algebraic Statistics, signal theory, coding theory and others.

For statistical inference, it is meaningful to know if a probability distribution, arising from a model, uniquely determines the parameters that produced it. When this happens, the parameters are called *identifiable*. The notion of *generic* identifiability for parametric models has been considered in [2] and in [16] §2.2. Indeed, conditions which guarantee the uniqueness of this decomposition, for generic tensors in the model, are quite important in the applications. When generic identifiability holds, the set of non-identifiable parameters has measure zero, thus parameter inference is still meaningful. Notice that many decomposition algorithms converge to *one* decomposition, hence a uniqueness result guarantees that the decomposition found is the chased one. We refer to [10] and its huge reference list, for more details.

Even from a purely theoretical point of view, the study of the decomposition shows some beautiful link between Multilinear Algebra and Projective Geometry.

*Corresponding author.

Email addresses: luca.chiantini@unisi.it (L. Chiantini), mll@unife.it (M. Mella), ottavian@math.unifi.it (G. Ottaviani)

The present paper is devoted to study one intriguing special case, which shows an exceptional behavior.

Among tensors $T \in \mathbb{C}^{a_1+1} \otimes \dots \otimes \mathbb{C}^{a_q+1}$ whose rank has the generic value, only one example is known when we have identifiability, that is $q = 3$, $a_1 = 1$, $a_2 = a_3 = a$. In this case we have the Kronecker-Weierstrass canonical form, see (10.3.1) in [12]. On the contrary, general tensors whose rank is *smaller* than the generic value, often have a unique decomposition.

Excluding the cases of matrices (tensors of order $q = 2$), identifiability is known to hold when the rank k is *small*. An evidence is given for $q = 3$ by the celebrated Kruskal's bound [11], which, for general tensors of given rank, is refined and extended in a series of papers (see Strassen's paper [15], the recent paper [7]).

Let's order the a_i 's so that $a_1 \leq \dots \leq a_q$. In [5] Corollary 8.4 it was proved that, with the assumption $a_q \geq \prod_{i=1}^{q-1} (a_i + 1) - \left(\sum_{i=1}^{q-1} a_i \right)$, the variety $\mathbb{P}^{a_1} \times \dots \times \mathbb{P}^{a_q}$ is k -identifiable if and only if

$$k \leq \prod_{i=1}^{q-1} (a_i + 1) - \left(1 + \sum_{i=1}^{q-1} a_i \right).$$

The general tensor of rank $k > \prod_{i=1}^{q-1} (a_i + 1) - \left(1 + \sum_{i=1}^{q-1} a_i \right)$ has not a unique decomposition. After this result, we say that a tensor is *unbalanced* if $a_q \geq \prod_{i=1}^{q-1} (a_i + 1) - \left(\sum_{i=1}^{q-1} a_i \right)$. It turns out that this range is one unity larger than the corresponding unbalanced range considered in [1] while studying the dimension of secant varieties of Segre varieties (see §8 of [5] where these two notions were compared). This phenomenon is quite common. The size of the tensors where we have not generic identifiability is often at the border of the range where we have defectivity for the corresponding secant varieties.

In the case $a_q \leq \prod_{i=1}^{q-1} (a_i + 1) - \left(1 + \sum_{i=1}^{q-1} a_i \right)$ the corresponding tensors are called *balanced*.

Only few examples of balanced tensors, whose rank is smaller than the generic value, are known to be not generically identifiable. We mention the case of tensors of rank 5 in $(\mathbb{C}^2)^{\otimes 5}$ ([4]) and, in dimension 3, tensors of rank 3 in $(\mathbb{C}^3)^{\otimes 3}$ (classical, see [15] §4) and tensors of rank 6 in $(\mathbb{C}^4)^{\otimes 3}$ ([7], Theorem 1.3).

A computer aided analysis (see [5] Theor. 7.5) shows that when the numbers a_i 's grow, sporadic examples disappear, and we expect that a general balanced tensor, of rank smaller than the generic value, is identifiable.

The present paper is devoted to illustrate one sporadic example, which we believe should be the last one, for balanced tensors of dimension $q = 3$. Namely, we use a geometric approach to show that general tensors of rank 8 in $\mathbb{C}^3 \otimes \mathbb{C}^6 \otimes \mathbb{C}^6$ are not uniquely decomposable. Notice that tensors of the mentioned type have generic rank equal to 9.

The proof of the non-uniqueness is based on the weak-defectivity principle, classically introduced by Terracini ([17]). We refer to [6, 13] and the introduction of [4] for an account of the geometric reduction of the problem.

In details, we prove that through 8 general points of the Segre variety $\mathbb{P}^2 \times \mathbb{P}^5 \times \mathbb{P}^5$,

which corresponds to simple tensors in $\mathbb{C}^3 \otimes \mathbb{C}^6 \otimes \mathbb{C}^6$, one can find a special fourfold Y which is the Segre-Veronese image of $\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$, embedded by forms of type $(3, 1, 1)$. Since through a general point of the span \mathbb{P}^{39} of Y one can find many linear 7-spaces which are 8-secant to Y , then by [6] Theorem 2.9, it follows the weak defectivity and the non identifiability of our tensors.

The example is interesting also because the subvariety Y , which produces the non-identifiability of tensors of rank 8 in $\mathbb{C}^3 \otimes \mathbb{C}^6 \otimes \mathbb{C}^6$, is quite complicate. In particular, we are unable to estimate how many 8-secant spaces to Y are there through a general point of the span \mathbb{P}^{39} . Consequently, we are unable to determine how many different decomposition are there, for a general tensor T as above. We simply know that the number is finite, and at least 6.

Let us mention that, from the geometrical point of view, the existence of the subvariety Y through 8 general points of $\mathbb{P}^2 \times \mathbb{P}^5 \times \mathbb{P}^5$ is proved by some "ad hoc" argument. A complete theory of special subvarieties that one can find through general points of Segre varieties, seems actually far beyond our reach.

2. Preliminaries

For basic facts about the geometric point of view on tensors we follow [12].

Given any irreducible projective variety X , we denote by $S_k(X)$ the k -th secant variety of X , that is the Zariski closure of the set $\bigcup_{x_1, \dots, x_k \in X} \langle x_1, \dots, x_k \rangle$. $S_k(X)$ is indeed the Zariski closure of the set of elements having X -rank equal to k .

In the space $\mathbb{P}^N = \mathbb{P}(\mathbb{C}^{a_1+1} \otimes \dots \otimes \mathbb{C}^{a_q+1})$, where $N = -1 + \prod_{i=1}^q (a_i + 1)$, the (projectification of the) cone X of simple tensors corresponds to the embedding of $\mathbb{P}^{a_1} \times \dots \times \mathbb{P}^{a_q}$, via the Segre map. The (projectification of the) cone of tensors of rank k is an open dense subset of the secant variety $S_k(X)$.

We recall from [7] def. 2.1 the following

Definition 1. X is called k -identifiable if the general element of $S_k(X)$ has a unique expression as the sum of k elements of X .

In our notation, we say that $\mathbb{P}^{a_1} \times \dots \times \mathbb{P}^{a_q}$ is k -identifiable if the general tensors in $\mathbb{C}^{a_1+1} \otimes \dots \otimes \mathbb{C}^{a_q+1}$ of rank k has a unique decomposition as a sum of simple tensors.

A complete list of known Segre varieties $X = \mathbb{P}(\mathbb{C}^{a_1+1}) \times \mathbb{P}(\mathbb{C}^{a_2+1}) \times \mathbb{P}(\mathbb{C}^{a_3+1})$, with $1 \leq a_1 \leq a_2 \leq a_3 \leq 6$, for which a computer based algorithm does not prove the k -identifiability, is provided in [7], §5, see also [5] §7. The list corresponds to the case of tensors of dimension 3, for which the algorithm cannot prove the uniqueness of the decomposition.

In all the examples, except for two of them, it is indeed well known that general tensors of rank k have infinitely many decompositions.

The two remaining cases are listed below:

(a_1, a_2, a_3)	k
<hr/>	
$(3, 3, 3)$	6
<hr/>	
$(2, 5, 5)$	8

In the first case, the effective proof that X is not 6-identifiable (and the general tensor of rank 6 has exactly 2 decompositions) is contained in [7], Theorem 1.3.

The latter case needs an "ad hoc" analysis which is the target of the present note.

Our main tool is to prove the existence of particular, very degenerate subvarieties Y , through k general points of the Segre variety $X = \mathbb{P}(\mathbb{C}^{a_1+1}) \times \mathbb{P}(\mathbb{C}^{a_2+1}) \times \mathbb{P}(\mathbb{C}^{a_3+1})$.

Indeed, we recall the following:

Theorem 1. *Let X be a projective, irreducible non-degenerate variety of dimension n in \mathbb{P}^r , $r > nk + k - 1$. Suppose that for any general k -tuples of points $x_1, \dots, x_k \in X$ one can find a subvariety Y of pure dimension $m > 0$ containing the points x_1, \dots, x_k , whose span has dimension*

$$\dim(\langle Y \rangle) = km + k - 1.$$

Assume that $S_k(Y) = \langle W \rangle$ and moreover assume that through a general point of $\langle Y \rangle$ one finds $\mu_k > 1$ k -secant $(k - 1)$ -linear spaces.

Then X is not k -identifiable. Indeed through a general point of $S_k(X)$ one finds at least μ_k k -secant $(k - 1)$ -linear spaces.

Proof. It is essentially Theorem 2.9 of [6].

3. Verifying the unidentifiability

From this point on, we focus our attention to the vector space V of tensors of type $\mathbb{C}^3 \otimes \mathbb{C}^6 \otimes \mathbb{C}^6$, which has dimension 108. From the projective point of view, simple tensors in V corresponds to points of the Segre embedding of $X = \mathbb{P}^2 \times \mathbb{P}^5 \times \mathbb{P}^5$ into $\mathbb{P}(\mathbb{C}^3 \otimes \mathbb{C}^6 \otimes \mathbb{C}^6) = \mathbb{P}^{107}$.

We also fix the rank $k = 8$, i.e. we consider the eighth secant variety $S_8(X)$. We know that X is not 8-defective, so that $S_8(X)$ has projective dimension 103 (see [7], §5). This means that the subvariety (cone) of tensors of rank 8 in V has the expected dimension 104.

By a computer-based calculation, using the M2 files available through the arXiv submission of [7], the guess is that X is 8-weakly defective, with a contact variety of dimension 4 and degree 108.

In order to verify the guess, we need a series of lemmas.

Lemma 1. *Fix eight general points P_1, \dots, P_8 of \mathbb{P}^5 and fix eight general points Q_1, \dots, Q_8 of \mathbb{P}^2 . Then there exists a Segre embedding $s: \mathbb{P}^2 \times \mathbb{P}^1 \rightarrow \mathbb{P}^5$ for which the line $s(\{Q_i\} \times \mathbb{P}^1)$ contains P_i for all i . In other words, each P_i lies in $s(\mathbb{P}^2 \times \mathbb{P}^1)$ and $\pi \circ s^{-1}(P_i) = Q_i$, where π is the projection $\mathbb{P}^2 \times \mathbb{P}^1 \rightarrow \mathbb{P}^2$.*

Proof. The embeddings $\mathbb{P}^2 \times \mathbb{P}^1 \rightarrow \mathbb{P}^5$ are parametrized by the quotient group $G = \text{Aut}(\mathbb{P}^5)/(\text{Aut}(\mathbb{P}^2) \times \text{Aut}(\mathbb{P}^1))$, which has dimension 24. Since the eight points P_i are general, we have an 8-dimensional family \mathcal{S} of embeddings s for which $P_1, \dots, P_8 \in s(\mathbb{P}^2 \times \mathbb{P}^1)$. Since the unique automorphism of \mathbb{P}^2 which fixes four general points is the identity, as s varies in \mathcal{S} , the family of 4-tuples $(\pi \circ s^{-1}(P_1), \dots, \pi \circ s^{-1}(P_4))$ dominates $(\mathbb{P}^2)^4$. Since the group $\text{Aut}(\mathbb{P}^2)$ acts transitively on the points $\pi \circ s^{-1}(P_5), \dots, \pi \circ s^{-1}(P_8)$, it follows that the orbit of the set $\{\pi \circ s^{-1}(P_5), \dots, \pi \circ s^{-1}(P_8)\}$, under $G \times \text{Aut}(\mathbb{P}^2)$, dominates $(\mathbb{P}^2)^8$. The claim follows.

Lemma 2. *Through 8 general points x_1, \dots, x_8 of X one can find a fourfold Y which corresponds to the embedding of $\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$ into \mathbb{P}^{39} , mapped by divisors of multidegree $(3, 1, 1)$.*

Proof. We send $\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$ to the three factors \mathbb{P}^2 , \mathbb{P}^5 and \mathbb{P}^5 , by using the identity on \mathbb{P}^2 and divisors $D = (1, 1, 0)$ and $D' = (1, 0, 1)$ respectively. Thus, we need to prove that we can arrange this map ζ so that the image passes through 8 general points of X .

The choice of eight general points in $\mathbb{P}^2 \times \mathbb{P}^5 \times \mathbb{P}^5$ corresponds to the choice of 8 general points in each factor. By the previous Lemma, for a general choice of points $Q_1, \dots, Q_8 \in \mathbb{P}^2$, $P_1, \dots, P_8 \in \mathbb{P}^5$ and $P'_1, \dots, P'_8 \in \mathbb{P}^5$, we can find divisors D, D' , which define Segre embeddings s, s' of $\mathbb{P}^2 \times \mathbb{P}^1$ into \mathbb{P}^5 , for which each P_i (resp. each P'_i) lies in the line $s(\{Q_i\} \times \mathbb{P}^1)$ (resp. $s'(\{Q_i\} \times \mathbb{P}^1)$).

It follows that $Y = \zeta(\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1)$ passes through each point x_i , $i = 1, \dots, 8$.

The following Lemma would be easy, provided one knows a table of 4-dimensional varieties in \mathbb{P}^{39} , whose 8-th secant order is different from 1. Since the table is missing, we need to compute directly what happens for the Segre product $\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$.

Lemma 3. *Let Y be an embedding of $\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$ into \mathbb{P}^{39} , through a divisor of type $(3, 1, 1)$. Then through a general point $y \in \mathbb{P}^{39}$ one can draw at least 6 spaces of dimension 7, which are 8-secant to Y .*

Proof. We consider the tangential projection from the tangent spaces at 7 general points $\{y_1, \dots, y_7\}$ of Y , which is a rational map $\tau_{y_1, \dots, y_7} \dashrightarrow Y \rightarrow \mathbb{P}^4$. By the Theorem 4.2 (vi) of [8] we have that the number of seven dimensional spaces which are 8-secant to Y and contain a general point $y \in \mathbb{P}^{39}$ is $\geq \deg \tau_{y_1, \dots, y_7}$, for a general choice of points $\{y_1, \dots, y_7\}$. So it is enough to show that $\deg \tau_{y_1, \dots, y_7} = 6$.

A computer based algorithm, implemented in M2 [9], which is available in the ancillary files of the arXiv submission of this paper, shows that there exists a 7-uple $\{\bar{y}_1, \dots, \bar{y}_7\}$ and a point $\bar{p} \in \mathbb{P}^4$ such that the fiber $\tau_{\bar{y}_1, \dots, \bar{y}_7}^{-1}(\bar{p})$ consists of 6 reduced points.

Consider the rational map

$$\tau : (Y)^7 \times \mathbb{P}^{39} \dashrightarrow (Y)^7 \times \mathbb{P}^4$$

induced by the tangential projection. That is

$$\tau(y_1, \dots, y_7, p) = (y_1, \dots, y_7, \tau_{y_1, \dots, y_7}(p)).$$

Then, after resolving the indeterminacy of the map τ , we get, from the Stein factorization, that the general fiber of τ consists of six points, so that $\deg \tau_{y_1, \dots, y_7} = 6$.

Remark 1. *Computer experiments, performed with the M2 file quoted in the above proof, show that the base locus of $\tau_{Y,s}$ consists of $2s$ lines for $s \leq 6$ (each tangent space at a point $y \in Y$ meets Y in two lines) and consists of 14 lines plus 4 extra points for $s = 7$. We do not know how to prove theoretically the existence of these 4 points in the base locus.*

Now we can use the approach of [6] to prove that X is not 8-identifiable.

Theorem 2. *X is not 8-identifiable. Through a general point $Q \in \mathbb{P}^{107}$ one can draw at least 6 spaces of dimension 7, which are 8-secant to X .*

Proof. Fix 8 general points $P_1, \dots, P_8 \in X$ and a general point $Q \in \langle P_1, \dots, P_8 \rangle$, so that Q is a general point of the 8-th secant variety of X . By Lemma 2, the eight points are contained in the image $Y \subset X$ of a Segre-Veronese embedding of $\mathbb{P}^2 \times \mathbb{P}^1 \times \mathbb{P}^1$ through a divisor of type $(3, 1, 1)$. Y spans a \mathbb{P}^{39} , which clearly contains Q , and Q is a general point of \mathbb{P}^{39} . By Lemma 3, one finds 6 linear spaces of (projective) dimension 7, which are 8-secant to Y and contain Q . Since these spaces are also 8-secant to X , the claim follows.

From a geometric point of view, Theorem 2.4 of [6] implies the following.

Corollary 1. *X is 8-weakly defective. A general hyperplane which is tangent to X at 8 general points, is also tangent along a subvariety Y of dimension 4, described above.*

Remark 2. *One would like to conclude that through a general point of the 8-secant variety of X one can find exactly 6 spaces of dimension 7, which are 8-secant to X .*

In other words, one would like to conclude that a general tensor of type $(3, 6, 6)$ and rank 8 can be written as a sum of 8 decomposable tensors in exactly 6 ways.

Unfortunately, we can only conclude that there are at least 6 decompositions. One reason is that the lower bound with $\deg \tau_{y_1, \dots, y_7}$ considered in the proof of Lemma 3 can be a strict inequality. For example, if Y is the 8-Veronese embedding of \mathbb{P}^2 , then Ranestad and Schreyer prove (see Theorem 1.7 (iv) of [14]) that a general polynomial of degree 8 has exactly 16 decompositions as the sum of 15 powers of linear forms. On the other hand, the tangential projection from 14 points has base locus given by the 14 points themselves and so its degree is $8^2 - 14 \cdot 4 = 8 < 16$.

Moreover, there could be more than one Segre-Veronese variety like Y , passing through 8 general points of X .

Using Terracini's interpretation of the secant varieties of Segre varieties ([17]), as explained in section 4 of [3], we can translate the main Theorem into a theorem on linear systems of matrices.

Corollary 2. *Let \mathcal{M} be a linear system of 5×5 matrices, with (affine) dimension 3. Assume that \mathcal{M} has rank 8, i.e. there are 8 matrices of rank 1 which generate all the elements of \mathcal{M} . Then there are at least 6 sets of 8 rank 1 matrices, whose spans contain \mathcal{M} .*

Proof. It is a straightforward consequence of the main Theorem of [3], see Remark 4.2 iii) there.

References

- [1] H Abo, G Ottaviani, and C Peterson. Induction for secant varieties of Segre varieties. *Trans. Amer. Math. Soc.*, 361(2):767–792, 2009.
- [2] E Allman, C Matias, and J Rhodes. Identifiability of parameters in latent structure models with many observed variables. *Annals of Statistics*, 37(6A):3009–3132, 2009.
- [3] E Ballico, A Bernardi, MV Catalisano, and L Chiantini. Grassman secants, identifiability, and linear systems of tensors. *Lin. Alg. Applic.*, 438:121–135, 2013.
- [4] C Bocci and L Chiantini. On the identifiability of binary Segre products. *J. Alg. Geom.*, 22:1–11, 2013.
- [5] C Bocci, L Chiantini, and G Ottaviani. Refined methods for the identifiability of tensors. *Annali di Matematica Pura e Applicata*, to appear, *arXiv:1303.6915*.
- [6] L Chiantini and C Ciliberto. On the k -th secant order of a projective variety. *J. London Math. Soc.*, 73(2):436–454, 2006.
- [7] L Chiantini and G Ottaviani. On generic identifiability of 3-tensors of small rank. *SIAM J. Matrix Anal. Appl.*, 33(3):1018–1037, 2012.
- [8] C Ciliberto and F Russo. Varieties with minimal secant degree and linear systems of maximal dimension on surfaces. *Adv. Math.*, 200(1):1–50, 2006.
- [9] D Grayson and M Stillman. Macaulay 2, a software system for research in algebraic geometry. www.math.uiuc.edu/Macaulay2/.
- [10] T Kolda and B Bader. Tensor Decompositions and Applications. *SIAM Review.*, 51(3):455–500, 2009.
- [11] JB Kruskal. Three-way arrays: rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics. *Linear Algebra Applic.*, 18(2):95–138, 1977.

- [12] JM Landsberg. *Tensors: Geometry and Applications.*, volume 128 of *Graduate Studies in Mathematics*. AMS, Providence, Rhode Island, 2012.
- [13] M Mella. Singularities of linear systems and the Waring problem. *Trans. Amer. Math. Soc.*, 358(12):5523–5538, 2006.
- [14] K Ranestad and FO Schreyer. Varieties of sums of powers. *J. Reine Angew. Math.*, 525:147–181, 2000.
- [15] V Strassen. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52:645–685, 1983.
- [16] S Sullivant and J Rhodes. Identifiability of large phylogenetic mixture models. *Bull. Math. Biol.*, 74(1):212–231, 2012.
- [17] A Terracini. Sulla rappresentazione delle coppie di forme ternarie mediante somme di potenze di forme lineari. *Ann. di Matem. Pura ed Appl.*, 24(III):91–100, 1915.